

PETERSON-GORENSTEIN-ZIERLER ALGORITHM FOR SKEW RS CODES

JOSÉ GÓMEZ-TORRECILLAS, F. J. LOBILLO, AND GABRIEL NAVARRO

ABSTRACT. We design a non-commutative version of the Peterson-Gorenstein-Zierler decoding algorithm for a class of codes that we call skew RS codes. These codes are left ideals of a quotient of a skew polynomial ring, which endow them of a sort of non-commutative cyclic structure. Since we work over an arbitrary field, our techniques may be applied both to linear block codes and convolutional codes. In particular, our decoding algorithm applies for block codes beyond the classical cyclic case.

1. INTRODUCTION

From a pure mathematical perspective, a linear block code is a vector subspace \mathcal{C} of \mathbb{F}^n , for some finite field \mathbb{F} . In order to get codes with good properties, it is usual to endow both \mathbb{F}^n and \mathcal{C} with additional algebraic structures. This is the case of BCH codes [1, 2], which become cyclic block codes, that is, ideals of the quotient ring $\mathbb{F}[x]/\langle x^n - 1 \rangle$ of the polynomial ring $\mathbb{F}[x]$ by the ideal generated by $x^n - 1$, see for instance [3]. Thus, they must be constructed by carefully selecting some factors of the polynomial $x^n - 1$. The reward is, for instance, that these codes are built with designed Hamming distance, and that there are several efficient decoding algorithms taking advantage of their rich algebraic structure. For example, the Peterson-Gorenstein-Zierler algorithm, that makes use of linear algebra techniques; the Sugiyama algorithm, which refines some steps by means of the polynomial arithmetic; or the Sudan-Guruswami algorithm for list decoding, which can be applied to the subclass of Reed-Solomon codes.

It is observed in [4, 5] that the number of potential “cyclic” codes \mathcal{C} of fixed length n is substantially increased if \mathcal{C} is required to be a left ideal of a suitable quotient ring of a skew polynomial ring $\mathbb{F}[x; \sigma]$, where σ is an automorphism of \mathbb{F} . It is worth to mention that, since [6, 7], non-commutative rings are used to endow convolutional codes with non trivial cyclic structures, see also [8, 9]. Recently, in [10], a non-commutative version of Sugiyama’s decoding algorithm [11] has been designed for convolutional codes built as certain left ideals of a simple ring of dimension n over the fraction field $\mathbb{F}(z)$ of the polynomial ring $\mathbb{F}[z]$, where z represents the delay operator. Concretely, see [12], SCCC codes are defined as left ideals of the ring $\mathbb{F}(z)[x; \sigma]/\langle x^n - 1 \rangle$, where σ is an \mathbb{F} -automorphism of order n of the field $\mathbb{F}(z)$.

In the present paper, a version of Peterson-Gorenstein-Zierler Algorithm for skew Reed-Solomon (RS) codes is proposed. Since we intend to cover both block and convolutional codes, we work with an abstract field L . Thus, we define a skew RS code as a left ideal, with generator carefully chosen, of a factor ring $L[x; \sigma]/\langle x^n - 1 \rangle$, where σ is an automorphism of L of (finite) order n . Of course, a verbatim translation of the original Peterson-Gorenstein-Zierler Algorithm [13, 14] makes no sense in our non-commutative context. Actually, even in the finite field block case, a skew cyclic code need not to be a cyclic code. Therefore our algorithm provides an efficient decoding procedure which applies beyond the classical Peterson-Gorenstein-Zierler algorithm. Our version shares with the block (commutative) one a general scheme where linear algebra tools and arguments play an important role, as, for instance, handling syndromes and computing errors. We also exploit the algebraic properties of the skew polynomial ring $L[x; \sigma]$ and of the field extension $L^\sigma \subset L$, which “encode” the skew cyclic structure.

The paper is structured as follows. Section 2 is devoted to state some technical results needed for proving the properties of skew RS codes and the steps of Algorithm 1. In particular, we would like to highlight what we have named the Circulant Lemma (Lemma 2.1) which is the key-tool that ensures the correctness of most of the methods developed in the paper. In Section 3 we define skew cyclic codes over a field and give a systematic procedure for generating them. Then we describe the subclass of skew RS codes and determine their Hamming distance. Section 4 is devoted to state and prove a Peterson-Gorenstein-Zierler algorithm for decoding skew RS codes. Finally, in Section 5, we provide a selection of examples aiming to illustrate the wide range of codes to which this algorithm can be applied.

2. SOME TECHNICAL RESULTS

Throughout this paper L will denote a field, $\sigma \in \text{Aut}(L)$ a field automorphism of order n , and L^σ , the invariant subfield. In this section we prove some technical results which will be used subsequently. We begin by what we call the *Circulant Lemma*, which is a particular case of [15, Corollary 4.13]. An

elementary proof can be found in [10]. We recall the reader that L is an L^σ -vector space of dimension n , the order of the automorphism, see e.g. [16, §4.5].

Lemma 2.1 (Circulant Lemma). [15, Corollary 4.13] *Let $\{\alpha_0, \dots, \alpha_{n-1}\}$ be an L^σ -basis of L . Then, for all $t \leq n$ and every subset $\{k_1, k_2, \dots, k_t\} \subseteq \{0, 1, \dots, n-1\}$,*

$$\begin{vmatrix} \alpha_{k_1} & \sigma(\alpha_{k_1}) & \dots & \sigma^{t-1}(\alpha_{k_1}) \\ \alpha_{k_2} & \sigma(\alpha_{k_2}) & \dots & \sigma^{t-1}(\alpha_{k_2}) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{k_t} & \sigma(\alpha_{k_t}) & \dots & \sigma^{t-1}(\alpha_{k_t}) \end{vmatrix} \neq 0.$$

Let us introduce the following notation, which will simplify some constructions in the sequel. For each $\gamma \in L$, we denote

$$\gamma^{[\sigma]} = \begin{pmatrix} \gamma \\ \sigma(\gamma) \\ \vdots \\ \sigma^{n-1}(\gamma) \end{pmatrix} \in L^n.$$

We also denote by $\varepsilon_i = (0, \dots, 1, \dots, 0)$, with $0 \leq i \leq n-1$, the vector with 1 in its i th position and 0, otherwise. These canonical vectors will be considered as rows or columns as required by the situation.

For the rest of this section, let us fix an L^σ -basis $\{\alpha_0, \dots, \alpha_{n-1}\}$ of L . A consequence of Lemma 2.1 is that any different vectors $\alpha_{k_1}^{[\sigma]}, \dots, \alpha_{k_t}^{[\sigma]}$ are linearly independent, where $\{k_1, \dots, k_t\} \subseteq \{0, \dots, n-1\}$.

Lemma 2.2. *Let $l, l_1, \dots, l_s \in \{0, \dots, n-1\}$. Then $\alpha_l^{[\sigma]}$ is a linear combination of $\alpha_{l_1}^{[\sigma]}, \dots, \alpha_{l_s}^{[\sigma]}$ if and only if $l = l_i$ for some $i \in \{1, \dots, s\}$.*

Proof. It follows trivially from Lemma 2.1. □

We now fix a set of indices $\{t_1, \dots, t_m\} \subseteq \{0, \dots, n-1\}$ and let

$$A = \left(\alpha_{t_1}^{[\sigma]} \mid \dots \mid \alpha_{t_m}^{[\sigma]} \right) = \begin{pmatrix} \alpha_{t_1} & \dots & \alpha_{t_m} \\ \sigma(\alpha_{t_1}) & \dots & \sigma(\alpha_{t_m}) \\ \vdots & \ddots & \vdots \\ \sigma^{n-1}(\alpha_{t_1}) & \dots & \sigma^{n-1}(\alpha_{t_m}) \end{pmatrix}.$$

By Lemma 2.1, $\text{rk } A = m$, where $\text{rk } A$ denotes the rank of A . Let $B \in \mathcal{M}_{m \times p}(L)$, a $m \times p$ matrix over L , such that $\text{rk } B = p$ and it has no zero row. Observe that then $p \leq m$ and $\text{rk } AB = p$. Let us split the set $\{1, \dots, m\}$ into two disjoint subsets G_1 and G_2 , where

$$\text{rk} \left(AB \mid \alpha_{t_i}^{[\sigma]} \right) = \begin{cases} p & \text{if } i \in G_1, \\ p+1 & \text{if } i \in G_2. \end{cases}$$

Denote by q the cardinal of the set G_1 .

Lemma 2.3. *Under the above conditions and notation, $q \leq p$. Moreover, $q = p$ if and only if $p = m$.*

Proof. Since $\text{rk } AB = p$, for each $i \in G_1$, $\alpha_{t_i}^{[\sigma]}$ is a linear combination of the columns of AB . Therefore, there exists $C \in \mathcal{M}_{p \times q}(L)$ such that the matrix formed by the column vectors $\alpha_{t_i}^{[\sigma]}$, for $i \in G_1$, equals ABC . Then $\text{rk } ABC = q$, by Lemma 2.1. Consequently, $q \leq p$.

Assume now $p = m$. Then B is non singular and $A = ABB^{-1}$. In particular, for each $1 \leq i \leq m$, $\alpha_{t_i}^{[\sigma]}$ is a linear combination of the columns of AB , i.e.

$$\text{rk} \left(AB \mid \alpha_{t_i}^{[\sigma]} \right) = m.$$

In other words, $G_1 = \{1, \dots, m\}$, so $q = m = p$.

Finally, suppose $q = p$ and assume $q < m$. We may reorder the columns of A in such a way that the first q columns correspond to the indices in G_1 , that is, $G_1 = \{1, \dots, q\}$ and $G_2 = \{q+1, \dots, m\}$. Let P be the permutation matrix which provides this reordering of the columns of A , hence $AB = APP^{-1}B$. Observe that $P^{-1}B$ also has rank p and all its rows are non zero. So, for simplicity, we abuse notation and denote also by A the matrix AP , and by B , the matrix $P^{-1}B$. We divide the matrices as

$$A = \left(A_0 \mid A_1 \right) \text{ and } B = \begin{pmatrix} B_0 \\ B_1 \end{pmatrix},$$

where A_0 encompasses the first q columns of A and B_0 , the first q rows of B . Then

$$q = \text{rk} \left(\begin{array}{c|c} AB & A_0 \end{array} \right) = \text{rk} \left(\begin{array}{c|c} A_0 B_0 + A_1 B_1 & A_0 \end{array} \right) = \text{rk} \left(\begin{array}{c|c} A_1 B_1 & A_0 \end{array} \right).$$

By hypothesis, the last row of B_1 is non zero. Let (b_{q+1}, \dots, b_m) be a column of B_1 with $b_m \neq 0$. Then

$$A_1 \begin{pmatrix} b_{q+1} \\ \vdots \\ b_m \end{pmatrix} = \alpha_{t_{q+1}}^{[\sigma]} b_{q+1} + \dots + \alpha_{t_m}^{[\sigma]} b_m$$

is a linear combination of the columns of A_0 , i.e. there exist $c_1, \dots, c_q \in L$ such that

$$\sum_{i=1}^q \alpha_{t_i}^{[\sigma]} c_i = \sum_{i=q+1}^m \alpha_{t_i}^{[\sigma]} b_i$$

and, consequently, $\alpha_{t_1}^{[\sigma]}, \dots, \alpha_{t_m}^{[\sigma]}$ are linearly dependent, contradicting Lemma 2.1. Thus $p = m$. \square

We shall also need the following straightforward technical lemma.

Lemma 2.4. *Let $H \in \mathcal{M}_{m \times n}(L)$ be a matrix in reduced row echelon form and let ε_i be a canonical vector of length n . Then*

$$\text{rk} \left(\frac{H}{\varepsilon_i} \right) = \text{rk } H$$

if and only if ε_i is a row of H .

3. SKEW CYCLIC CODES

In this section we introduce skew cyclic codes over L and give some of their properties. Let us denote by R the skew polynomial ring $L[x; \sigma]$, see [17]. Given $f, g \in R$, we say that f right divides g , $f \mid_r g$, if $Rg \subseteq Rf$, i.e. the remainder of the left division of g by f is zero. For $\gamma \in L$, the right evaluation of f in γ is the remainder of the left division of f by $x - \gamma$. As R is a left (and right) PID, there are least common multiples and greatest common divisors of polynomials on both sides. We use the notation

$$[f, g]_\ell \text{ and } (f, g)_r$$

to refer to the least common left multiple and the greatest common right divisor of a pair $f, g \in R$. Then,

$$Rf + Rg = R(f, g)_r \text{ and } Rf \cap Rg = R[f, g]_\ell.$$

Since we are assuming that the order of σ is n , the polynomial $x^n - 1$ is central in R , so we may consider the quotient ring $\mathcal{R} = L[x; \sigma] / \langle x^n - 1 \rangle$. Throughout, we shall see the elements in \mathcal{R} as polynomials of degree lower than n . As an L -vector space, \mathcal{R} is isomorphic to L^n via the coordinate map $\mathbf{v} : \mathcal{R} \rightarrow L^n$, mapping each polynomial (of degree lower than n) to the vector formed by its coefficients. We shall use freely this identification all along the paper.

Linear codes over L are vector subspaces of L^n . When $L = \mathbb{F}_q$, a finite field, these are block codes, whilst, when $L = \mathbb{F}_q(z)$, these are convolutional codes [18]. So we adopt the coding theory terminology over L . Following this philosophy, we may define the *weight* of a vector and the *Hamming distance* of a code as usual. All these notions are going to be used without further mention in the framework of a vector space over L .

Definition 1. A *skew cyclic code* over L is a vector subspace $\mathcal{C} \leq L^n$ such that $\mathbf{v}^{-1}(\mathcal{C})$ is a left ideal of \mathcal{R} . Equivalently, it is a vector subspace $\mathcal{C} \leq L^n$ such that

$$(a_0, \dots, a_{n-2}, a_{n-1}) \in \mathcal{C} \Rightarrow (\sigma(a_{n-1}), \sigma(a_0), \dots, \sigma(a_{n-2})) \in \mathcal{C}.$$

These codes have been introduced under the name of σ -cyclic codes in [4], when L is a finite field, and skew cyclic convolutional codes in [12], when $L = \mathbb{F}_q(z)$.

Any left ideal of \mathcal{R} is principal, so every skew cyclic code is generated by a polynomial in R . Obviously, the generator can be taken as a right divisor of $x^n - 1$, so it is important to determine the decompositions of $x^n - 1$ in R . First we recall some notions about skew polynomials. Let $\gamma \in L$, the i th-norm of γ is defined to be

$$N_i(\gamma) = \gamma \sigma(\gamma) \dots \sigma^{i-1}(\gamma).$$

Norms are useful to evaluate skew polynomials. In fact, if $f = \sum_{i \geq 0} f_i x^i \in R$, the remainder of the left division of f by $x - \gamma$ is

$$(1) \quad \sum_{i \geq 0} f_i N_i(\gamma).$$

This is an easy computation that can be found in [19] in a more general context. We also recall some formulas in order to ease the reading of the paper. Concretely, if $\alpha, \beta, \gamma \in L$ such that $\beta = \alpha^{-1}\sigma(\alpha)$, then

$$(2) \quad \begin{aligned} N_i(\sigma^k(\gamma)) &= \sigma^k(N_i(\gamma)), \\ N_i(\sigma^k(\beta)) &= \sigma^k(\alpha)^{-1}\sigma^{k+i}(\alpha). \end{aligned}$$

We shall follow the systematic method described in [12] in order to construct skew cyclic codes. By the Normal Basis Theorem, we may choose an element $\alpha \in L$ such that $\{\alpha, \sigma(\alpha), \dots, \sigma^{n-1}(\alpha)\}$ is a basis of L as an L^σ -vector space. Set now $\beta = \alpha^{-1}\sigma(\alpha)$. We shall fix this notation for the rest of the paper.

Lemma 3.1. *For any subset $T = \{t_1, t_2, \dots, t_m\} \subseteq \{0, 1, \dots, n-1\}$, the polynomial*

$$g = [x - \sigma^{t_1}(\beta), x - \sigma^{t_2}(\beta), \dots, x - \sigma^{t_m}(\beta)]_\ell$$

has degree m . Consequently, if $x - \sigma^s(\beta) \mid_r g$, then $s \in T$.

Proof. The statement can be proved following the same steps of the proof of [10, Lemma 2] □

As a consequence of Lemma 3.1,

$$x^n - 1 = [x - \beta, x - \sigma(\beta), \dots, x - \sigma^{n-1}(\beta)]_\ell,$$

since, by (2), $N_n(\sigma^k(\beta)) = \sigma^k(\alpha)^{-1}\sigma^{k+n}(\alpha) = \sigma^k(\alpha^{-1}\alpha) = 1$, and then $x - \sigma^k(\beta)$ right divides $x^n - 1$ for all $0 \leq k \leq n-1$. Therefore, given $\{t_1, \dots, t_k\} \subseteq \{0, 1, \dots, n-1\}$, the polynomial $g = [x - \sigma^{t_1}(\beta), \dots, x - \sigma^{t_k}(\beta)]_\ell$ generates a left ideal $\mathcal{R}g$ such that $\mathfrak{v}(\mathcal{R}g)$ is a skew cyclic code of dimension $n - k$.

For the convenience, we call β -roots to the elements of the set $\{\beta, \sigma(\beta), \dots, \sigma^{n-1}(\beta)\}$. By (1) and (2), given a polynomial $f = \sum_{i=0}^{n-1} f_i x^i \in \mathcal{R}$,

$$(3) \quad x - \sigma^j(\beta) \mid_r f \iff \sum_{i=0}^{n-1} f_i N_i(\sigma^j(\beta)) = 0 \iff \sum_{i=0}^{n-1} f_i \sigma^{i+j}(\alpha) = 0.$$

Let then N be the matrix formed by the norms of the β -roots,

$$(4) \quad N = \begin{pmatrix} N_0(\beta) & N_0(\sigma(\beta)) & \dots & N_0(\sigma^{n-1}(\beta)) \\ N_1(\beta) & N_1(\sigma(\beta)) & \dots & N_1(\sigma^{n-1}(\beta)) \\ \vdots & \vdots & \ddots & \vdots \\ N_{n-1}(\beta) & N_{n-1}(\sigma(\beta)) & \dots & N_{n-1}(\sigma^{n-1}(\beta)) \end{pmatrix},$$

the components of $\mathfrak{v}(f)N = (f_0, \dots, f_{n-1})N$ are the right evaluations of f in the set of β -roots, i.e. the vector formed by the left remainders of f by the polynomials $x - \sigma^i(\beta)$ for $i = 0, \dots, n-1$. Hence, the diagram

$$\begin{array}{ccc} \mathcal{R} & & \\ \downarrow \mathfrak{v} & \searrow ev & \\ L^n & \xrightarrow{N} & L^n, \end{array}$$

is a commutative diagram of L -linear isomorphisms, where ev maps each polynomial f to the n -tuple formed by the left remainders of f by $x - \sigma^i(\beta)$ for $i = 0, \dots, n-1$. Indeed, by Lemma 2.1, N is non singular, so it provides a change of basis. We call the *set of β -roots of f* to the set formed by the β -roots γ verifying that $x - \gamma \mid_r f$, that is by those corresponding to the zero coordinates of $(f_0, \dots, f_{n-1})N$.

We say that, a non-constant right divisor $f \mid_r x^n - 1$, *fully β -decomposes* if there exists $\{t_1, \dots, t_m\} \subseteq \{0, 1, \dots, n-1\}$ such that

$$f = [x - \sigma^{t_1}(\beta), \dots, x - \sigma^{t_m}(\beta)]_\ell.$$

Observe that, by Lemma 3.1, $\deg f = m$, the cardinal of the set of β -roots of f .

Lemma 3.2. *Let $f = \sum_{i=0}^m f_i x^i \in \mathcal{R}$ with $f_m \neq 0$ and*

$$M_f = \begin{pmatrix} f_0 & f_1 & \dots & f_m & 0 & \dots & 0 \\ 0 & \sigma(f_0) & \dots & \sigma(f_{m-1}) & \sigma(f_m) & \dots & 0 \\ & & \ddots & & & \ddots & \\ 0 & \dots & 0 & \sigma^{n-m-1}(f_0) & \dots & \dots & \sigma^{n-m-1}(f_m) \end{pmatrix}_{(n-m) \times n}.$$

Then the rows of M_f are a basis of $\mathfrak{v}(\mathcal{R}f)$ as an L -vector space. Moreover, f fully β -decomposes if and only if

$$\text{rref}(M_f N) = \left(\begin{array}{c} \varepsilon_{i_1} \\ \vdots \\ \varepsilon_{i_{n-m}} \end{array} \right)$$

for some $0 \leq i_1 < \dots < i_{n-m} \leq n-1$, where rref denotes the reduced row echelon form.

Proof. An L -basis of $\mathcal{R}f$ is $\{f, xf, \dots, x^{n-m-1}f\}$, whose coordinates correspond to the rows of M_f . Now, $f = [x - \sigma^{t_1}(\beta), \dots, x - \sigma^{t_m}(\beta)]_\ell$ if and only if any left multiple of f is also a left multiple of $x - \sigma^{t_i}(\beta)$ for $1 \leq i \leq m$, if and only if the t_i -th columns of $M_f N$ are zero for $i = 1, \dots, m$. Since $M_f N$ has $n-m$ rows, rank $n-m$ and $n-m$ non zero columns, the result follows. \square

We shall need the following result.

Lemma 3.3. *Let $f, g \in \mathcal{R}$ be fully β -decomposable polynomials. Then $(f, g)_r$ and $[f, g]_\ell$ are also fully β -decomposable.*

Proof. Since f, g are fully β -decomposable, there exist subsets $T_1, T_2 \subseteq \{0, \dots, n-1\}$ such that

$$f = [\{x - \sigma^i(\beta)\}_{i \in T_1}]_\ell \text{ and } g = [\{x - \sigma^i(\beta)\}_{i \in T_2}]_\ell.$$

Hence, it is straightforward that

$$[f, g]_\ell = [\{x - \sigma^i(\beta)\}_{i \in T_1 \cup T_2}]_\ell.$$

On the other hand,

$$[\{x - \sigma^i(\beta)\}_{i \in T_1 \cap T_2}]_\ell \mid_r (f, g)_r.$$

The formula $\deg f + \deg g = \deg (f, g)_r + \deg [f, g]_\ell$ and Lemma 3.1 give the equality. \square

We now may define the class of skew cyclic codes for which the decoding algorithm of the next section can be applied.

Definition 2. Under the conditions and notation of this section, a *skew Reed-Solomon (RS) code* of designed Hamming distance δ is a skew cyclic code \mathcal{C} such that $\mathfrak{v}^{-1}(\mathcal{C})$ is generated by a polynomial $[x - \sigma^r(\beta), x - \sigma^{r+1}(\beta), \dots, x - \sigma^{r+\delta-2}(\beta)]_\ell$ for some $r \geq 0$.

Theorem 3.4. *A skew RS code of designed Hamming distance δ has Hamming distance δ . Consequently, it is an MDS code.*

Proof. The result can be proved analogously to the proof of [10, Theorem 4]. \square

4. A PETERSON-GORENSTEIN-ZIERLER DECODING ALGORITHM

Throughout this section \mathcal{C} denotes a skew RS code as described in Definition 2. Without loss of generality, we may assume that \mathcal{C} is a narrow-sense skew RS code, i.e. we may set $r = 0$. This is because we always may write $\alpha' = \sigma^r(\alpha)$, which also provides a normal basis, and then $\sigma^r(\beta) = \beta' = (\alpha')^{-1}\sigma(\alpha')$, so $[x - \beta', \dots, x - \sigma^{\delta-2}(\beta')]_\ell$ is a generator of \mathcal{C} . Therefore, suppose that the left ideal $\mathfrak{v}^{-1}(\mathcal{C})$ is generated by $g = [x - \beta, x - \sigma(\beta), \dots, x - \sigma^{\delta-2}(\beta)]_\ell$ for some $2 \leq \delta \leq n$. By Theorem 3.4, the Hamming distance of \mathcal{C} is exactly δ and, following Algorithm 1 below, it can correct up to $t = \lfloor \frac{\delta-1}{2} \rfloor$ errors, the error correction capability of the code.

Suppose now that a message $m = (m_0, \dots, m_{n-\delta})$ must be transmitted through a noisy channel. Since we are identifying \mathcal{C} with $\mathfrak{v}^{-1}(\mathcal{C})$, $m = \sum_{i=0}^{n-\delta} m_i x^i$. The message m is encoded to a codeword $c = mg$ and $y = c + e$ is received, where $e = e_1 x^{k_1} + \dots + e_\nu x^{k_\nu}$, with $\nu \leq t$, is the error polynomial. The purpose of this section is to develop an algorithm, following the scheme of the classical Peterson-Gorenstein-Zierler decoding algorithm, for computing this error.

For each $0 \leq i \leq 2t-1$, the i th *syndrome* s_i of the received polynomial y is defined to be the left remainder of y by $x - \sigma^i(\beta)$. Since c is right divisible by $x - \sigma^i(\beta)$ for $i = 0, \dots, \delta-2$; it follows, by (2), that

$$\begin{aligned} s_i &= \sum_{j=0}^{n-1} y_j N_j(\sigma^i(\beta)) = \sum_{j=1}^{\nu} e_j N_{k_j}(\sigma^i(\beta)) \\ (5) \quad &= \sum_{j=1}^{\nu} e_j \sigma^i(\alpha^{-1}) \sigma^{i+k_j}(\alpha) = \sigma^i(\alpha^{-1}) \sum_{j=1}^{\nu} e_j \sigma^{i+k_j}(\alpha). \end{aligned}$$

Proposition 4.1. *The error values (e_1, \dots, e_ν) are the unique solution of the linear system*

$$X \begin{pmatrix} \sigma^{k_1}(\alpha) & \sigma^{k_1+1}(\alpha) & \dots & \sigma^{k_1+\nu-1}(\alpha) \\ \sigma^{k_2}(\alpha) & \sigma^{k_2+1}(\alpha) & \dots & \sigma^{k_2+\nu-1}(\alpha) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{k_\nu}(\alpha) & \sigma^{k_\nu+1}(\alpha) & \dots & \sigma^{k_\nu+\nu-1}(\alpha) \end{pmatrix} = (\alpha s_0, \sigma(\alpha)s_1, \dots, \sigma^{\nu-1}(\alpha)s_{\nu-1}).$$

Proof. Observe that

$$\begin{pmatrix} \sigma^{k_1}(\alpha) & \sigma^{k_1+1}(\alpha) & \dots & \sigma^{k_1+\nu-1}(\alpha) \\ \sigma^{k_2}(\alpha) & \sigma^{k_2+1}(\alpha) & \dots & \sigma^{k_2+\nu-1}(\alpha) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{k_\nu}(\alpha) & \sigma^{k_\nu+1}(\alpha) & \dots & \sigma^{k_\nu+\nu-1}(\alpha) \end{pmatrix} = \begin{pmatrix} \sigma^{k_1}(\alpha) & \sigma(\sigma^{k_1}(\alpha)) & \dots & \sigma^{\nu-1}(\sigma^{k_1}(\alpha)) \\ \sigma^{k_2}(\alpha) & \sigma(\sigma^{k_2}(\alpha)) & \dots & \sigma^{\nu-1}(\sigma^{k_2}(\alpha)) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{k_\nu}(\alpha) & \sigma(\sigma^{k_\nu}(\alpha)) & \dots & \sigma^{\nu-1}(\sigma^{k_\nu}(\alpha)) \end{pmatrix},$$

which has non zero determinant, by Lemma 2.1. By (5)

$$\sigma^i(\alpha)s_i = \sum_{j=1}^{\nu} e_j \sigma^{i+k_j}(\alpha),$$

so the result follows. \square

By Proposition 4.1, the decoding process is therefore reduced to find the error positions $\{k_1, \dots, k_\nu\}$. We define the *error locator* polynomial as

$$\lambda = [x - \sigma^{k_1}(\beta), x - \sigma^{k_2}(\beta), \dots, x - \sigma^{k_\nu}(\beta)]_\ell.$$

By Lemma 3.1, λ has degree ν , and, once λ is known, the error positions can be determined.

Recall that $f = \sum_{k=0}^{n-1} f_k x^k \in \mathcal{R}\lambda$ if and only if $x - \sigma^{k_j}(\beta) \mid_r f$ for all $j = 1, \dots, \nu$; or, equivalently, $\sum_{k=0}^{n-1} f_k N_k(\sigma^{k_j}(\beta)) = 0$ for all $j = 1, \dots, \nu$. Therefore, $(f_0, \dots, f_{n-1}) \in \mathfrak{v}(\mathcal{R}\lambda)$ if and only if f satisfies the equation $(f_0, \dots, f_{n-1})T = 0$, where

$$T = \begin{pmatrix} N_0(\sigma^{k_1}(\beta)) & N_0(\sigma^{k_2}(\beta)) & \dots & N_0(\sigma^{k_\nu}(\beta)) \\ N_1(\sigma^{k_1}(\beta)) & N_1(\sigma^{k_2}(\beta)) & \dots & N_1(\sigma^{k_\nu}(\beta)) \\ \vdots & \vdots & \ddots & \vdots \\ N_{n-1}(\sigma^{k_1}(\beta)) & N_{n-1}(\sigma^{k_2}(\beta)) & \dots & N_{n-1}(\sigma^{k_\nu}(\beta)) \end{pmatrix}.$$

Now, by (2), $N_k(\sigma^{k_j}(\beta)) = \sigma^{k_j}(\alpha)^{-1} \sigma^{k_j+k}(\alpha)$, then $\mathcal{R}\lambda$ corresponds to the left kernel of the matrix

$$\Sigma = \begin{pmatrix} \sigma^{k_1}(\alpha) & \sigma^{k_2}(\alpha) & \dots & \sigma^{k_\nu}(\alpha) \\ \sigma^{k_1+1}(\alpha) & \sigma^{k_2+1}(\alpha) & \dots & \sigma^{k_\nu+1}(\alpha) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{k_1+n-1}(\alpha) & \sigma^{k_2+n-1}(\alpha) & \dots & \sigma^{k_\nu+n-1}(\alpha) \end{pmatrix} = \begin{pmatrix} \Sigma_0 \\ \Sigma_1 \end{pmatrix},$$

where Σ_0 comprises the first $\nu + 1$ rows of Σ . Let us consider the matrix

$$E = \begin{pmatrix} e_1 & \sigma^{-1}(e_1) & \dots & \sigma^{-\nu+1}(e_1) \\ e_2 & \sigma^{-1}(e_2) & \dots & \sigma^{-\nu+1}(e_2) \\ \vdots & \vdots & \ddots & \vdots \\ e_\nu & \sigma^{-1}(e_\nu) & \dots & \sigma^{-\nu+1}(e_\nu) \end{pmatrix}.$$

Thus $S = \Sigma E$ is an $(n \times \nu)$ -matrix whose (k, i) -component is given by $\sum_{j=1}^{\nu} \sigma^{-i}(e_j) \sigma^{k_j+k}(\alpha)$. Observe that, by (5), whenever $k + i < 2t - 1$, this component can be written as $\sigma^{-i}(s_{k+i}) \sigma^k(\alpha)$, so that we may divide $S = \begin{pmatrix} S_0 \\ S_1 \end{pmatrix}$, where

$$S_0 = \begin{pmatrix} s_0 \alpha & \sigma^{-1}(s_1) \alpha & \dots & \sigma^{-\nu+1}(s_{\nu-1}) \alpha \\ s_1 \sigma(\alpha) & \sigma^{-1}(s_2) \sigma(\alpha) & \dots & \sigma^{-\nu+1}(s_\nu) \sigma(\alpha) \\ \vdots & \vdots & \ddots & \vdots \\ s_\nu \sigma^\nu(\alpha) & \sigma^{-1}(s_{\nu+1}) \sigma^\nu(\alpha) & \dots & \sigma^{-\nu+1}(s_{2\nu-1}) \sigma^\nu(\alpha) \end{pmatrix}_{(\nu+1) \times \nu},$$

whose coefficients can be computed from the received polynomial y . In order to calculate the parameter ν , the number of error positions, we may follow the scheme of Peterson-Gorenstein-Zierler algorithm for

BCH block codes, see e.g. [3, §5.4]. For any $1 \leq r \leq t$, let us denote by S^r the matrix

$$S^r = \begin{pmatrix} s_0\alpha & \sigma^{-1}(s_1)\alpha & \cdots & \sigma^{-r+1}(s_{r-1})\alpha \\ s_1\sigma(\alpha) & \sigma^{-1}(s_2)\sigma(\alpha) & \cdots & \sigma^{-r+1}(s_r)\sigma(\alpha) \\ \vdots & \vdots & \ddots & \vdots \\ s_t\sigma^t(\alpha) & \sigma^{-1}(s_{t+1})\sigma^t(\alpha) & \cdots & \sigma^{-r+1}(s_{t+r-1})\sigma^t(\alpha) \end{pmatrix}_{(t+1) \times r}.$$

Observe that, for all $r \leq t$, $S^r = \Sigma^t E^r$, where

$$E^r = \begin{pmatrix} e_1 & \sigma^{-1}(e_1) & \cdots & \sigma^{-r+1}(e_1) \\ e_2 & \sigma^{-1}(e_2) & \cdots & \sigma^{-r+1}(e_2) \\ \vdots & \vdots & \ddots & \vdots \\ e_\nu & \sigma^{-1}(e_\nu) & \cdots & \sigma^{-r+1}(e_\nu) \end{pmatrix}_{\nu \times r}$$

and

$$\Sigma^t = \begin{pmatrix} \sigma^{k_1}(\alpha) & \sigma^{k_2}(\alpha) & \cdots & \sigma^{k_\nu}(\alpha) \\ \sigma^{k_1+1}(\alpha) & \sigma^{k_2+1}(\alpha) & \cdots & \sigma^{k_\nu+1}(\alpha) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{k_1+t}(\alpha) & \sigma^{k_2+t}(\alpha) & \cdots & \sigma^{k_\nu+t}(\alpha) \end{pmatrix}_{(t+1) \times \nu}.$$

Lemma 4.2. *For each $r \leq t$, $\text{rk } S^r = \text{rk } \Sigma E^r = \text{rk } E^r$.*

Proof. By Lemma 2.1, $\text{rk } \Sigma = \text{rk } \Sigma^t = \nu$. Using Sylvester's rank inequality,

$$\min\{\text{rk } \Sigma, \text{rk } E^r\} \geq \text{rk } \Sigma E^r \geq \text{rk } \Sigma + \text{rk } E^r - \nu = \text{rk } E^r.$$

Then $\text{rk } \Sigma E^r = \text{rk } E^r$. Analogously, $\text{rk } S^r = \text{rk } E^r$. \square

Hence, we may calculate the greatest r such that S^r has full rank. By Lemma 4.2, it is also the greatest integer $r \leq t$ such that E^r and ΣE^r have full rank. We shall denote by μ such a maximum.

Lemma 4.3. *For each $\mu \leq r \leq t$, $\mu = \text{rk } E^r = \text{rk } S^r$. Consequently, $\mu \leq \nu$.*

Proof. By Lemma 4.2, $\mu = \text{rk } E^\mu = \text{rk } S^\mu$, so assume $\mu < r$. By maximality of μ , the $(\mu+1)$ th column of E^r is a linear combination of the μ preceding columns. Applying σ^{-1} we get that the $(\mu+2)$ th column is a linear combination of the columns at positions from the second to the $(\mu+1)$ th, and hence a linear combination of the first μ columns. Repeating the process we obtain that all columns from the $(\mu+1)$ th to r th are linear combinations of the first μ columns, which implies that $\text{rk } E^r = \mu$. Since E^r has ν rows, $\mu \leq \nu$. Finally, $\text{rk } S^r = \mu$ again by Lemma 4.2. \square

Proposition 4.4. *The left kernel V of the matrix ΣE^μ is a skew cyclic code. Consequently, $\mathbf{v}^{-1}(V) = \mathcal{R}\rho$ for some polynomial $\rho \in \mathcal{R}$ of degree μ . Moreover, ρ is a right divisor of λ .*

Proof. The first statement is ensured by proving that, if $(a_0, \dots, a_{n-2}, a_{n-1}) \in V \subseteq L^n$, then we have $(\sigma(a_{n-1}), \sigma(a_0), \dots, \sigma(a_{n-2})) \in V$. This is due to the fact that $xa_i x^i = \sigma(a_i) x^{i+1}$ for every $0 \leq i \leq n-1$. Suppose then $(a_0, a_1, \dots, a_{n-1}) \Sigma E^\mu = 0$. The maximality of μ ensures that the last column of $E^{\mu+1}$ is a linear combination of the former μ columns. Hence $(a_0, a_1, \dots, a_{n-1}) \Sigma E^{\mu+1} = 0$. Therefore,

$$\begin{aligned} 0 &= (a_0, a_1, \dots, a_{n-1}) \Sigma E^{\mu+1} \\ &= (a_0, a_1, \dots, a_{n-1}) \left(\begin{array}{c|c} 0 & I_{n-1} \\ \hline 1 & 0 \end{array} \right) \left(\begin{array}{c|c} 0 & 1 \\ \hline I_{n-1} & 0 \end{array} \right) \Sigma E^{\mu+1} \\ &= (a_{n-1}, a_0, \dots, a_{n-2}) \left(\begin{array}{c|c} 0 & 1 \\ \hline I_{n-1} & 0 \end{array} \right) \Sigma E^{\mu+1}. \end{aligned}$$

Applying σ to this matrix equation (componentwise),

$$(\sigma(a_{n-1}), \sigma(a_0), \dots, \sigma(a_{n-2})) \left(\begin{array}{c|c} 0 & 1 \\ \hline I_{n-1} & 0 \end{array} \right) \sigma(\Sigma) \sigma(E^{\mu+1}) = 0.$$

Observe that $\Sigma = \left(\begin{array}{c|c} 0 & 1 \\ \hline I_{n-1} & 0 \end{array} \right) \sigma(\Sigma)$ and $\sigma(E^{\mu+1}) = \left(\begin{array}{c|c} \sigma(e_1) & \\ \vdots & E^\mu \\ \sigma(e_\nu) & \end{array} \right)$, so

$$(\sigma(a_{n-1}), \sigma(a_0), \dots, \sigma(a_{n-2})) \Sigma \left(\begin{array}{c|c} \sigma(e_1) & \\ \vdots & E^\mu \\ \sigma(e_\nu) & \end{array} \right) = 0.$$

In particular, $(\sigma(a_{n-1}), \sigma(a_0), \dots, \sigma(a_{n-2}))\Sigma E^\mu = 0$, so $(\sigma(a_{n-1}), \sigma(a_0), \dots, \sigma(a_{n-2})) \in V$, as desired. Now, any left ideal of \mathcal{R} is principal, so $\mathfrak{v}^{-1}(V)$ is generated by a polynomial $\rho \in \mathcal{R}$. Since $\mathfrak{v}(\mathcal{R}\lambda)$ is the left kernel of the matrix Σ , it follows that $\mathcal{R}\lambda \subseteq \mathcal{R}\rho$, hence ρ right divides λ . Finally, the dimension of $\mathcal{R}\rho$ as an L -vector space is $n - \deg \rho$. By Lemma 4.2, $\text{rk } \Sigma E^\mu = \mu$, so $\deg \rho = \mu$. \square

Lemma 4.5. *The reduced column echelon form of S^t is*

$$\text{rcef}(S^t) = \left(\begin{array}{c|c} \frac{I_\mu}{a_0 \cdots a_{\mu-1}} & 0_{(t+1) \times (t-\mu)} \\ \hline & H' \end{array} \right),$$

where I_μ is the $\mu \times \mu$ identity matrix and $a_0, \dots, a_{\mu-1} \in L$ such that $\rho = x^\mu - \sum_{i=0}^{\mu-1} a_i x^i$.

Proof. By Lemma 4.3, $\text{rk } S^t = \mu = \text{rk } S^\mu$, so

$$\text{rcef}(S^t) = \left(\text{rcef}(S^\mu) \mid 0_{(t+1) \times (t-\mu)} \right).$$

Recall that S^μ consists of the first $t+1$ rows of ΣE^μ and both have the same rank μ , therefore $\text{rcef}(S^\mu)$ is composed by the first $t+1$ rows of $\text{rcef}(\Sigma E^\mu)$. By Proposition 4.4, $\mathfrak{v}(\mathcal{R}\rho)$ is the left kernel of the matrix $\text{rcef}(\Sigma E^\mu)$. A non zero solution of the homogeneous system

$$(6) \quad X \left(\text{rcef}(\Sigma E^\mu) \mid \frac{0}{I_{n-(\mu+1)}} \right) = 0$$

is a non zero element of $\mathfrak{v}(\mathcal{R}\rho)$ whose $n - (\mu+1)$ last coordinates are zero. Since ρ has degree μ , and its degree is minimal in $\mathcal{R}\rho$, it follows that $\mathfrak{v}(\rho)$ is the unique solution, up to scalar multiplication, of (6). Let S_0^μ be formed by the first $\mu+1$ rows of S^μ . Then

$$\text{rcef}(S^\mu) = \left(\frac{\text{rcef}(S_0^\mu)}{H'} \right).$$

Further column reductions using the identity matrix in the right block of the matrix in (6), allow us to see that ρ is also the non zero solution, up to scalar multiplication, of the homogeneous system

$$(7) \quad X \left(\frac{\text{rcef}(S_0^\mu)}{0} \mid \frac{0}{I_{n-(\mu+1)}} \right) = 0.$$

The size of $\text{rcef}(S_0^\mu)$ is $(\mu+1) \times \mu$. Moreover $\text{rk } \text{rcef}(S_0^\mu) = \mu$ because the space of solutions of (7) has dimension 1. So there is only one row of $\text{rcef}(S_0^\mu)$ without a pivot. If this row is not the last row then there would be a non zero polynomial in $\mathcal{R}\rho$ of degree strictly below μ , which is impossible. Hence

$$\text{rcef}(S_0^\mu) = \left(\frac{I_\mu}{a_0 \cdots a_{\mu-1}} \right).$$

Finally, $(-a_0, \dots, -a_{\mu-1}, 1, 0, \dots, 0)$ is a non zero solution of (7), it follows $\rho = x^\mu - \sum_{i=0}^{\mu-1} a_i x^i$. \square

Lemma 4.6. *If the left ideal $\mathcal{R}\rho$ corresponds, via \mathfrak{v} , to the left kernel of a matrix H , then $H = \Sigma B$ for some $B \in \mathcal{M}_{\nu \times \mu}(L)$ which has no zero row.*

Proof. The statement comes from the commutative diagram of L -vector spaces

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathcal{R}\rho & \longrightarrow & \mathcal{R} & \xrightarrow{\cdot H} & \mathcal{R}/\mathcal{R}\rho \longrightarrow 0 \\ & & \uparrow & & \parallel & & \uparrow \cdot B \\ 0 & \longrightarrow & \mathcal{R}\lambda & \longrightarrow & \mathcal{R} & \xrightarrow{\cdot \Sigma} & \mathcal{R}/\mathcal{R}\lambda \longrightarrow 0 \end{array}$$

If $\mathcal{R}\rho$ corresponds to the left kernel of a matrix H , there exists a surjective L -linear map $\mathcal{R}/\mathcal{R}\lambda \rightarrow \mathcal{R}/\mathcal{R}\rho$ defined by right multiplication by a $(\nu \times \mu)$ -matrix B , such that $\Sigma B = H$. Since $\mathcal{R}\rho$ is also the left kernel of ΣE^μ , there exists a non singular $(\mu \times \mu)$ -matrix P such that $\Sigma E^\mu P = \Sigma B$. Since Σ defines a surjective linear map, then $E^\mu P = B$. Finally, B is obtained from E^μ by elementary operations on the columns. Since E^μ has no zero row, B is so. \square

There is a strong connection between ρ and λ , since the error locator polynomial is minimal for ρ in the following sense.

Proposition 4.7. *Let $\lambda' \in \mathcal{R}$ be a fully β -decomposable polynomial which is a left multiple of ρ . Then $\lambda \mid_r \lambda'$.*

Proof. By Proposition 4.4, $\rho \mid_r \lambda$ and, by hypothesis, $\rho \mid_r \lambda'$. Then $\rho \mid_r (\lambda, \lambda')_r$. Moreover, by Lemma 3.3, $(\lambda, \lambda')_r$ is fully β -decomposable. Let us denote $\phi = (\lambda, \lambda')_r$. We claim that $\phi = \lambda$, which implies the statement.

Indeed, $\mathcal{R}\lambda \subseteq \mathcal{R}\phi$, then $\mathcal{R}\phi$ corresponds to the left kernel of ΣQ , where Q is a full rank matrix. Analogously, $\mathcal{R}\phi \subseteq \mathcal{R}\rho$, so there exists Q' of full rank such that $\mathcal{R}\rho$ is the left kernel of $\Sigma Q Q'$. By Lemma 4.6, $\Sigma Q Q' = \Sigma B$, where B has full rank and no zero row. Hence, $Q Q' = B$, because Σ defines a surjective linear map, and Q has no zero row.

Since $\phi \mid_r \lambda$, any β -root of ϕ must be a β -root of λ , so that it belongs to the set $\{\sigma^{k_1}(\beta), \dots, \sigma^{k_\nu}(\beta)\}$. Observe that, by (3), $\sigma^{k_j}(\beta)$ is a β -root of ϕ if and only if

$$\text{rk} \left(\Sigma Q \mid \sigma^{k_j}(\alpha)^{[\sigma]} \right) = \text{rk} \Sigma Q.$$

Hence, by Lemma 2.3, ϕ being fully β -decomposable implies $\{\sigma^{k_1}(\beta), \dots, \sigma^{k_\nu}(\beta)\}$ is the set of β -roots of ϕ . Thus $\phi = \lambda$. \square

At this point we have developed all the ingredients needed to compute the error locator polynomial, which completes the steps for designing the Peterson-Gorenstein-Zierler algorithm for skew cyclic codes, see Algorithm 1.

Algorithm 1 PGZ decoding algorithm

Input: A received transmission $y = (y_0, \dots, y_{n-1}) \in L^n$ with no more than t errors.

Output: The error $e = (e_0, \dots, e_{n-1})$ such that $y - e \in \mathcal{C}$

- 1: **for** $0 \leq i \leq 2t - 1$ **do**
 - 2: $s_i \leftarrow \sum_{j=0}^{n-1} y_j N_j(\sigma^i(\beta))$
 - 3: **if** $s_i = 0$ for all $0 \leq i \leq 2t - 1$ **then**
 - 4: **return** 0.
 - 5: $S^t \leftarrow (\sigma^{-j}(s_{i+j})\sigma^i(\alpha))_{0 \leq i \leq t, 0 \leq j \leq t-1}$
 - 6: Compute
$$\text{rcef}(S^t) = \left(\frac{I_\mu}{\frac{a_0 \cdots a_{\mu-1}}{H'}} \mid 0_{(t+1) \times (t-\mu)} \right).$$
 - 7: $\rho = (\rho_0, \dots, \rho_\mu) \leftarrow (-a_0, \dots, -a_{\mu-1}, 1)$ and $\rho_N \leftarrow (\rho_0, \dots, \rho_\mu, 0, \dots, 0)N$.
 - 8: $\{k_1, \dots, k_\nu\} \leftarrow$ zero coordinates of ρ_N
 - 9: **if** $\mu \neq \nu$ **then**
 - 10: $M_\rho \leftarrow \begin{pmatrix} \rho_0 & \rho_1 & \cdots & \rho_\mu & 0 & \cdots & 0 \\ 0 & \sigma(\rho_0) & \cdots & \sigma(\rho_{\mu-1}) & \sigma(\rho_\mu) & \cdots & 0 \\ & & \ddots & & & \ddots & \\ 0 & \cdots & 0 & \sigma^{n-\mu-1}(\rho_0) & \cdots & \cdots & \sigma^{n-\mu-1}(\rho_\mu) \end{pmatrix}_{(n-\mu) \times n}$
 - 11: $N_\rho \leftarrow M_\rho N$
 - 12: $H_\rho \leftarrow \text{rref}(N_\rho)$
 - 13: $H' \leftarrow$ matrix obtained removing all rows of H_ρ different from ε_i for any i .
 - 14: $\{k_1, \dots, k_\nu\} \leftarrow$ zero column coordinates of H'
 - 15: Solve the linear system $(x_1, \dots, x_\nu) (\Sigma^{\nu-1})^\top = (\alpha s_0, \sigma(\alpha)s_1, \dots, \sigma^{\nu-1}(\alpha)s_{\nu-1})$
 - 16: **return** (e_0, \dots, e_{n-1}) with $e_i = x_i$ for $i \in \{k_1, \dots, k_\nu\}$, and zero otherwise.
-

Theorem 4.8. Let L be a field, $\sigma \in \text{Aut}(L)$ of order n and L^σ the invariant subfield. Let the set $\{\alpha, \sigma(\alpha), \dots, \sigma^{n-1}(\alpha)\}$ be a normal basis of L over L^σ and $\beta = \alpha^{-1}\sigma(\alpha)$. Let $\mathcal{R} = L[x; \sigma]/\langle x^n - 1 \rangle$, $g = [x - \beta, \dots, x - \sigma^{\delta-2}(\beta)]_\ell$ and \mathcal{C} the skew RS code such that $\mathbf{v}^{-1}(\mathcal{C}) = \mathcal{R}g$. Then Algorithm 1 correctly finds the error $e = (e_0, \dots, e_{n-1})$ of any received vector if the number of non zero coordinates of e is $\nu \leq t = \lfloor \frac{\delta-1}{2} \rfloor$.

Proof. After the initial settings, Line 7 computes a right divisor $\rho = \sum_{i=0}^{\mu} \rho_i x^i$ of the error locator λ by Proposition 4.4 and Lemma 4.5.

By (3), Line 8 computes all the β -roots of ρ . By Lemma 3.1, $\nu = \mu$ if and only if ρ is fully β -decomposable. In this case, by Proposition 4.7, $\rho = \lambda$.

If $\nu \neq \mu$, since $\deg \rho = \mu$, the rows of M_ρ generate $\mathcal{R}\rho$ as an L -vector space, and the rows of N_ρ also generates $\mathcal{R}\rho$ under the change of basis corresponding to N . Since H_ρ is the reduced row echelon form of M_ρ , then its rows are also a basis of $\mathcal{R}\rho$ as an L -vector space. By Lemma 3.2, the rows of H' generate

an L -vector subspace $\mathcal{R}\lambda'$ for some fully β -decomposable polynomial λ' . Since H' is obtained removing some rows of H_ρ , it follows that $\rho \mid_r \lambda'$.

Let us now prove that $\lambda' = \lambda$. By Proposition 4.7, $\lambda \mid_r \lambda'$. Suppose that $\lambda \neq \lambda'$, then the matrix $H_\lambda = \text{rref}(M_\lambda N)$ contains an additional row ε_d not in H' . Since $\mathcal{R}\lambda \subseteq \mathcal{R}\rho$, $\text{rk}\left(\frac{H_\rho}{\varepsilon_d}\right) = \text{rk}(H_\rho)$. By Lemma 2.4, ε_d is a row of H_ρ , so it is not removed in Line 13 of Algorithm 1. Hence ε_d belongs to H' , a contradiction. Thus $\lambda = \lambda'$ and the error positions are computed. By Proposition 4.1, Line 15 computes the error values. \square

Remark 1. The complexity of Algorithm 1 is dominated by the computation of the reduced echelon forms in Lines 6 and 12. The theoretical efficiency of these lines belong to $\mathcal{O}(t^3)$ and $\mathcal{O}(n^3)$, respectively. Since, in the worst case, $t \approx n/2$, the complexity of Algorithm 1 is in $\mathcal{O}(n^3)$. Nevertheless, in most cases, the **if** part (Lines 10-14) is not executed. Indeed, μ is not the true number of errors ν if and only if the determinant

$$|E^\nu| = \begin{vmatrix} e_1 & \sigma^{-1}(e_1) & \cdots & \sigma^{-\nu+1}(e_1) \\ e_2 & \sigma^{-1}(e_2) & \cdots & \sigma^{-\nu+1}(e_2) \\ \vdots & \vdots & \ddots & \vdots \\ e_\nu & \sigma^{-1}(e_\nu) & \cdots & \sigma^{-\nu+1}(e_\nu) \end{vmatrix} = 0$$

Therefore, taking coordinates with respect to a fixed basis of L over L^σ , we deduce that the set of errors $\{e_1, \dots, e_\nu\}$ which yields $|E^\nu| = 0$ is contained in the determinantal algebraic subvariety of $(L^\sigma)^{\nu n}$ determined by the common zeroes of all $\nu \times \nu$ minors. The dimension of this variety is known to be at most $n - \nu + 1$ (see, e.g., [20, Exercise 10.10]), which is strictly smaller than $\nu n = \dim_{L^\sigma} L^\nu$ if $\nu > 1$.

5. DECODING SKEW RS CODES

In this section we illustrate the scope of application of Algorithm 1 with some examples. The calculations have been made with the aid of the mathematical software Sagemath [21].

5.1. σ -Cyclic codes. Our algorithm can be applied to some of the σ -cyclic codes defined in [4]. These are defined as left ideals of the factor algebra $\mathbb{F}_q[x; \sigma]/\langle x^n - 1 \rangle$, where \mathbb{F}_q is the finite field of q elements and the order of σ divides n . So Algorithm 1 can be applied whenever n is the order of σ , as the following example shows.

Example 1. Let $\mathbb{F} = \mathbb{F}_2(a)$ be the field with 2^{12} elements, where $a^{12} + a^7 + a^6 + a^5 + a^3 + a + 1 = 0$. Except for 0 and 1, we write the elements of \mathbb{F} as powers of a . Consider $\sigma : \mathbb{F} \rightarrow \mathbb{F}$ defined by $\sigma = \tau^{10}$, where τ is the Frobenius automorphism, i.e. $\sigma(a) = a^{1024}$. The order of σ is 6, so a skew cyclic code over \mathbb{F} is a left ideal of the quotient algebra $\mathcal{R} = \mathbb{F}[x; \sigma]/\langle x^6 - 1 \rangle$. Let now take $\alpha = a$, which provides a normal basis of \mathbb{F} as an \mathbb{F}^σ -vector space, and $\beta = \sigma(a)a^{-1} = a^{1023}$. Then, the images of β under the powers of σ give us the set $\{a^{1023}, a^{3327}, a^{3903}, a^{4047}, a^{4083}, a^{4092}\}$. We may also construct the matrix of the change of basis

$$N = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ a^{1023} & a^{3327} & a^{3903} & a^{4047} & a^{4083} & a^{4092} \\ a^{255} & a^{3135} & a^{3855} & a^{4035} & a^{4080} & a^{1020} \\ a^{63} & a^{3087} & a^{3843} & a^{4032} & a^{1008} & a^{252} \\ a^{15} & a^{3075} & a^{3840} & a^{960} & a^{240} & a^{60} \\ a^3 & a^{3072} & a^{768} & a^{192} & a^{48} & a^{12} \end{pmatrix}.$$

Let us consider the skew RS code generated by

$$g = [x - a^{1023}, x - a^{3327}, x - a^{3903}, x - a^{4047}]_\ell = x^4 + a^{2103}x^3 + a^{687}x^2 + a^{1848}x + a^{759}.$$

By Theorem 3.4, it has Hamming distance 5, so, following Algorithm 1, it can correct until 2 errors. The reader may check easily that this code is not cyclic in the usual sense. Suppose then we need to send the message $m = x + a$, so the encoded polynomial to be transmitted is $c = mg = x^5 + a^{3953}x^4 + a^{1333}x^3 + a^{2604}x^2 + a^{1596}x + a^{760}$. After the transmission, we receive a polynomial $y = x^5 + a^{3953}x^4 + a^{671}x^3 + a^{2604}x^2 + a^{1596}x + a^{3699}$, i.e. there are two errors at positions 0 and 3. Actually, we have added the error $e = a^2 + a^3x^3$.

We first calculate the full matrix of syndromes

$$\begin{pmatrix} a^{3170} & a^{2390} \\ a^{2645} & a^{428} \\ a^{107} & a^{248} \end{pmatrix},$$

and its reduced column echelon form

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ a^{1950} & a^{3315} \end{pmatrix}.$$

Therefore, the rank of the matrix is two, and the monic polynomial in the left kernel is $\rho = x^2 + a^{3315}x + a^{1950}$, and $\rho N = (0, a^{210}, a^{2685}, 0, a^{1155}, a^{3945})$. So ρ is the error locator polynomial and the error positions are 0 and 3. Now, in order to compute the error values, we need to solve the system

$$(e_0, e_3) \begin{pmatrix} a & a^{1024} \\ a^{64} & a^{16} \end{pmatrix} = (a^{3170}, a^{2645}),$$

which yields $e_0 = a^2$ and $e_3 = a^3$, as expected.

Let us now suppose that we receive a polynomial $y = x^5 + a^{3953}x^4 + a^{671}x^3 + a^{2604}x^2 + a^{1596}x + a^{3699}$, that is, we have added the error $e = a^2 + a^{1367}x^3$. In this case, the syndrome matrix and its reduced column echelon form are

$$\begin{pmatrix} a^{59} & a^{65} \\ a^{1040} & a^{1046} \\ a^{2309} & a^{2315} \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ a^{981} & 0 \\ a^{2250} & 0 \end{pmatrix},$$

respectively. Then $\rho = x + a^{981}$, and $\rho N = (a^{1437}, a^{1281}, a^{4053}, a^9, a^{3149}, a^{3853})$. Therefore, ρ is not the error locator polynomial. We then compute the matrices

$$M_\rho = \begin{pmatrix} a^{981} & 1 & 0 & 0 & 0 & 0 \\ 0 & a^{1269} & 1 & 0 & 0 & 0 \\ 0 & 0 & a^{1341} & 1 & 0 & 0 \\ 0 & 0 & 0 & a^{1359} & 1 & 0 \\ 0 & 0 & 0 & 0 & a^{3411} & 1 \end{pmatrix}$$

and

$$N_\rho = M_\rho N = \begin{pmatrix} a^{1437} & a^{1281} & a^{4053} & a^9 & a^{3149} & a^{3853} \\ a^{2406} & a^{576} & a^{1845} & a^{978} & a^{1799} & a^{1984} \\ a^{3672} & a^{3471} & a^{1293} & a^{2244} & a^{3509} & a^{493} \\ a^{1941} & a^{3171} & a^{1155} & a^{513} & a^{1889} & a^{1144} \\ a^{2532} & a^{3096} & a^{3168} & a^{1104} & a^{1484} & a^{283} \end{pmatrix}.$$

Now, the reduced row echelon form of N_ρ is as follows

$$H_\rho = \begin{pmatrix} 1 & 0 & 0 & a^{2667} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

If we remove the first row, we find zero columns at positions 0 and 3, i.e. the error positions. Finally, we solve the system

$$(e_0, e_3) \begin{pmatrix} a & a^{1024} \\ a^{64} & a^{16} \end{pmatrix} = (a^{59}, a^{1040}),$$

which yields $e_0 = a^2$ and $e_3 = a^{1367}$.

5.2. Skew cyclic convolutional codes. In [12] it is taken into consideration a novel approach to cyclicity for convolutional codes by introducing the so-called skew cyclic convolutional codes (SCCCs). This perspective considers the embedding of a polynomial ring $\mathbb{F}[z]$ into its field of fractions $\mathbb{F}(z)$, so that SCCCs are skew cyclic codes over $\mathbb{F}(z)$.

Example 2. Let $\mathbb{F} = \mathbb{F}_4 = \{0, 1, a, a^2\}$ be the field with four elements, $\mathbb{F}(z)$ the field of rational functions over \mathbb{F} and $\sigma : \mathbb{F}(z) \rightarrow \mathbb{F}(z)$ the automorphism of order 5 defined by $\sigma(z) = (z + a)/(z + a^2)$. The working ring is $\mathcal{R} = \mathbb{F}(z)[x; \sigma]/(x^5 - 1)$. Let us consider $\alpha = z \in \mathbb{F}(z)$, $\beta = \alpha^{-1}\sigma(\alpha) = (z + a)/(z^2 + a^2z)$, and the skew cyclic code generated by $g = [x - \beta, x - \sigma(\beta), x - \sigma^2(\beta), x - \sigma^3(\beta)]_\ell$. Concretely,

$$g = x^4 + \left(\frac{z+a}{z^5+a^2z}\right)x^3 + \left(\frac{az^5+a^2z^4+az+a^2}{z^5+a^2z^4+a^2z+a}\right)x^2 + \left(\frac{a^2z^5+z^4+z+a}{z^4+a^2}\right)x + \frac{az^5+a^2z^4}{a^2z^5+a^2z^4+az+a}.$$

So it can correct up to two errors. Suppose that g is transmitted and we receive the polynomial y given by

$$x^4 + \left(\frac{1}{z^4+a^2}\right)x^3 + \left(\frac{az^5+a^2z^4+az+a^2}{z^5+a^2z^4+a^2z+a}\right)x^2 + \left(\frac{a^2z^6+z^5+z^2+az+1}{z^5+a^2z}\right)x + \frac{az^5+a^2z^4}{a^2z^5+a^2z^4+az+a},$$

i.e., $y=g+e$, where $e = a/(z^5 + a^2z)x^3 + 1/(z^5 + a^2z)x$. We follow Algorithm 1 and compute the matrix of syndromes,

$$\begin{pmatrix} \frac{a^2z^2+az+a^2}{a^2z^7+a^2z^6+a^2z^5+a^2z^3+az^2+az} & \frac{az^7+a^2z^6+a^2z^5+az^4+az^3+a^2z^2+a^2z+a}{z^3+az^2+az+a^2} \\ \frac{z^2+z+a^2}{az^7+a^2z^6+a^2z^5+a^2z^3+az^2+az} & \frac{az^6+az^5+z^4+az^2+az+1}{az^2+z} \\ \frac{a^2z^2+z+a^2}{az^6+a^2z^5+z^2+az} & \frac{az^7+z^6+z^5+az^4+az^3+z^2+z+a}{a^2z^2+a^2z+a^2} \end{pmatrix}$$

and its reduced column echelon form

$$\begin{pmatrix} 1 & 0 \\ \frac{a^2z^4+az^2+z+a}{z^4+az^3+az^2+z} & 0 \\ \frac{az^3+az^2+1}{z^2+a^2z+1} & 0 \end{pmatrix}.$$

Therefore,

$$\rho = x + \frac{a^2z^4+az^2+z+a}{z^4+az^3+az^2+z} \equiv \left(\frac{a^2z^4+az^2+z+a}{z^4+az^3+az^2+z}, 1, 0, 0, 0 \right).$$

In this case, the matrix N is given as follows:

$$N = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ \frac{z+a}{z^2+a^2z} & \frac{a^2z^2+z+a}{az^2+a^2z} & \frac{z}{z^2+a^2z+a} & \frac{z^2+z+1}{a^2z+a^2} & \frac{z^2+z}{a^2z+a} \\ \frac{az+a}{z^2} & \frac{a^2z+a}{az^2+1} & \frac{z^2+a^2z}{a^2z^2+a^2} & a^2z^2+z & \frac{z^2+a^2z+a}{a^2z^2+1} \\ \frac{a^2}{az^2+a^2z} & \frac{a^2z^2+1}{az^2+a^2z+a} & \frac{z^2}{az+a} & \frac{a^2z^2+a}{z+a^2} & \frac{a^2z^2+a^2}{a^2z^2+a^2} \\ \frac{a^2z+a}{z^2+z} & \frac{a^2z^2+az}{az^2+1} & \frac{z^2+az}{az^2+a^2z+1} & \frac{az^2+z+a^2}{az} & \frac{a^2z+a^2}{z^2+z+1} \end{pmatrix}.$$

Now, ρN has no zero coordinate, so ρ is not the error locator polynomial. Following Algorithm 1,

$$M_\rho = \begin{pmatrix} \frac{a^2z^4+az^2+z+a}{z^4+az^3+az^2+z} & 1 & 0 & 0 & 0 \\ 0 & \frac{az^4+z^3+z^2+az}{a^2z^3+az^2+a^2z+a^2} & 1 & 0 & 0 \\ 0 & 0 & \frac{a^2z^3+az^2+a}{z^4+z^2+a^2z+a^2} & 1 & 0 \\ 0 & 0 & 0 & \frac{a^2z^4+a^2z^3+a^2z^2+az+1}{a^2z^3+a^2z^2+z} & 1 \end{pmatrix}.$$

Hence, the reduced row echelon form of $M_\rho N$ is as follows:

$$H_\rho = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{az^2+1}{z+a^2} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

If we remove the second row, the resultant matrix has a zero column at positions 1 and 3. Finally, we may find the error values by solving the linear system

$$(e_1, e_3) \begin{pmatrix} \frac{z+a}{z+a^2} & \frac{az+a}{z+1} \\ \frac{a}{z+a} & \frac{a^2z+a}{z+1} \end{pmatrix} = \left(\frac{a^2z^2+az+a^2}{a^2z^7+a^2z^6+a^2z^5+az^3+az^2+az}, \frac{z^2+z+a^2}{az^7+az^6+z^3+z^2} \right),$$

which yields $e_1 = 1/(z^5 + a^2z)$ and $e_3 = a/(z^5 + a^2z)$.

5.3. Skew cyclic codes over a cyclotomic field. Here we show an additional example of a class of skew cyclic codes over a non-conventional field. The base field of this kind of codes is a cyclotomic field $\mathbb{Q}(\chi)$, where χ is an n th root of unity.

Example 3. Let $L = \mathbb{Q}(\chi)$, where χ is a primitive 7th root of unit, and $\sigma : L \rightarrow L$ defined by $\sigma(\chi) = \chi^3$. In this case, the order of σ is 6. Let us set $\alpha = \chi$, $\beta = \alpha^{-1}\sigma(\alpha) = \chi^2$ and $\delta = 5$, so that the corresponding skew RS code is generated by

$$g = 2x^4 + (-\chi^5 - \chi^3 - \chi^2)x^3 + (\chi^3 + \chi + 1)x^2 + (\chi^5 + \chi^4 + 1)x + \chi^5 - \chi^2 + \chi + 1.$$

Suppose that g is transmitted, i.e. the message $m = 1$ is sent, and we receive the polynomial

$$y = 2x^4 + (-\chi^5 - \chi^3 - \chi^2)x^3 + (\chi^3 + 2\chi + 1)x^2 + (\chi^5 + \chi^4 + 1)x + \chi^5 - \chi^2 + \chi + 1.$$

Since $t = 2$, the syndrome matrix S^t , and its reduced column echelon form, are given by

$$\begin{pmatrix} \chi^3 & 1 \\ 1 & \chi^4 \\ \chi^5 & \chi^2 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ \chi^4 & 0 \\ \chi^2 & 0 \end{pmatrix},$$

respectively. Therefore $\mu = 1$ and $\rho = x - \chi^4$. Now, the matrix of norms is

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ \chi^2 & b & \chi^4 & \chi^5 & \chi & \chi^3 \\ \chi & \chi^3 & \chi^2 & b & \chi^4 & \chi^5 \\ \chi^5 & \chi & \chi^3 & \chi^2 & b & \chi^4 \\ \chi^3 & \chi^2 & b & \chi^4 & \chi^5 & \chi \\ \chi^4 & \chi^5 & \chi & \chi^3 & \chi^2 & b \end{pmatrix},$$

where $b = -\chi^5 - \chi^4 - \chi^3 - \chi^2 - \chi - 1$, so

$$\rho_N = (-\chi^4 + \chi^2, -\chi^5 - 2\chi^4 - \chi^3 - \chi^2 - \chi - 1, 0, \chi^5 - \chi^4, -\chi^4 + \chi, -\chi^4 + \chi^3).$$

By Algorithm 1, there is a single error at position 2 whose value may be computed by solving the equation $e_1\chi^2 = \chi^3$. That is, the error is $e = \chi x^2$.

FUNDING

Research supported by grants MTM2016-78364-P, MTM2013-41992-P and TIN2013-41990-R from Ministerio de Economía y Competitividad and from Fondo Europeo de Desarrollo Regional FEDER.

REFERENCES

- [1] Hocquenghem A. Codes correcteurs d'erreurs. Chiffres 1959;2:147–156.
- [2] Bose R, Ray-Chaudhuri D. On a class of error correcting binary group codes. Inform. Control 1960;3(1):68–79.
- [3] Huffman W, Pless V. Fundamentals of Error-Correcting Codes. Cambridge: University Press; 2010.
- [4] Boucher D, Geiselmann W, Ulmer F. Skew-cyclic codes. Appl. Algebr. Eng. Comm 2007;18:379–389.
- [5] Boucher D, Ulmer F. Coding with skew polynomial rings. J. Symb. Comp. 2009;44(12):1644–1656.
- [6] Piret P. On a class of alternating cyclic convolutional codes. IEEE T. Inform. Theory 1975;21(1):64–69.
- [7] Piret P. Structure and constructions of cyclic convolutional codes. IEEE T. Inform. Theory 1976;22(2):147–155.
- [8] Ross C. On the structure of convolutional and cyclic convolutional codes. IEEE T. Inform. Theory 1979;25(6):676–683.
- [9] Gluesing-Luerssen H, Schmale W. On cyclic convolutional codes. Acta Appl. Math. 2004;82(2):183–237.
- [10] Gómez-Torrecillas J, Lobillo FJ, Navarro G. A Sugiyama-like decoding algorithm for convolutional codes. arXiv:1607.07187
- [11] Sugiyama Y, Kasahara M, Hirasawa S, Namekawa T. A method for solving key equation for decoding goppa codes. Inform. Control 1975;27(1):87–99.
- [12] Gómez-Torrecillas J, Lobillo FJ, Navarro G. A new perspective of cyclicity in convolutional codes. IEEE T. Inform. Theory 2016;62(5):2702–2706.
- [13] Peterson WW. Encoding and error-correction procedures for the bose-chaudhuri codes. IRE T. Inform. Theory 1960;6(4):459–470.
- [14] Gorenstein D, Zierler N. A class of error-correcting codes in p^m symbols. J. Soc. Ind. Appl. Math. 1961;9(2):207–214.
- [15] Lam TY, Leroy A. Vandermonde and wronskian matrices over division rings, J. Algebra 1988;119(2):308–336.
- [16] Jacobson N. Basic Algebra I. 2nd ed. New York: W. H. Freeman and Company; 1985.
- [17] Ore O. Theory of non-commutative polynomials. Ann. Math. 1933;34(3):480–508.
- [18] Forney Jr GD. Convolutional codes I: algebraic structure. IEEE T. Inform. Theory 1970;16(6):720–738.
- [19] Leroy A. Pseudo linear transformations and evaluation in Ore extensions., B. Belg. Math. Soc-Sim 1995;2:321–347.
- [20] Eisenbud D. Commutative Algebra with a View Toward Algebraic Geometry. Vol. 150. Graduate Texts in Mathematics. New York: Springer; 1995.
- [21] The Sage Developers. SageMath, the Sage Mathematics Software System (Version 7.2). <http://www.sagemath.org>.2016.

E-mail address: gomezj@ugr.es

CITIC AND DEPARTMENT OF ALGEBRA, UNIVERSITY OF GRANADA, SPAIN

E-mail address: jlobillo@ugr.es

CITIC AND DEPARTMENT OF ALGEBRA, UNIVERSITY OF GRANADA, SPAIN

E-mail address: gnavarro@ugr.es

CITIC AND DEPARTMENT OF COMPUTER SCIENCES AND AI, UNIVERSITY OF GRANADA, SPAIN